

SVEUČILIŠTE U ZAGREBU
FILOZOFSKI FAKULTET
ODSJEK ZA INFORMACIJSKE I KOMUNIKACIJSKE ZNANOSTI
Ak. god. 2018./ 2019.

Zrinka Jagečić

Alati za poslovnu inteligenciju

Završni rad

Mentor: doc.dr.sc. Vedran Juričić

Zagreb, lipanj, 2019.

Izjava o akademskoj čestitosti

Izjavljujem i svojim potpisom potvrđujem da je ovaj rad rezultat mog vlastitog rada koji se temelji na istraživanjima te objavljenom i citiranoj literaturi. Izjavljujem da nijedan dio rada nije napisan na nedozvoljen način, odnosno da je prepisan iz necitiranog rada, te da nijedan dio rada ne krši bilo čija autorska prava. Također izjavljujem da nijedan dio rada nije korišten za bilo koji drugi rad u bilo kojoj drugoj visokoškolskoj, znanstvenoj ili obrazovnoj ustanovi.

(potpis)

Sadržaj

Sadržaj.....	iv
1. Uvod.....	1
2. Poslovna inteligencija.....	2
2.1. Razvoj poslovne inteligencije	3
2.1.1. Postignute promjene.....	3
2.1.2. Mogući problemi	4
3. Skladište podataka.....	7
3.1. Dimenzijski model podataka.....	9
3.1.1. Kocke.....	10
3.1.2. Činjenična tablica	10
3.1.3. Dimenzijska tablica.....	12
3.1.4. Zvezdasti model.....	13
3.1.5. Pahuljasti model	14
3.1.6. SQL	15
4. Microsoft alati	18
4.1. Microsoft Excel.....	18
4.1.1. Pivot tablica	20
4.1.2. Power Pivot	21
4.2. Microsoft SQL Server	22
4.2.1. SQL Server Database Engine	24
4.2.2. SQL Server Reporting Services.....	24
4.2.3. SQL Server Integration Services	24
4.2.4. SQL Server Analysis Services.....	26
4.3. Power BI	26
4.4. Usporedba alata.....	31

5. Zaključak.....	32
6. Literatura.....	33
7. Popis slika.....	34
Sažetak	35
Summary.....	36

1. Uvod

U ovom radu će se prikazati primjer i praksa alata koji se može koristiti u procesu poslovne inteligencije. Prvo je potrebno uopće razraditi sami pojam poslovne inteligencije (engl. *Business Intelligence*), odnosno, BI, skraćeno od punog engleskog naziva. Poslovna inteligencija se danas postavlja kao neizbježan faktor uspjeha poslovanja neke organizacije. Bitno je spomenuti koju ulogu ona ima u modernom poslovanju te zbog čega se većina upravo odlučuje za neku verziju vizualizacije vlastitog poslovanja. Objasniti će se odakle i kako je došlo do ideje za potrebom poslovne inteligencije u svakodnevici te navesti njene prednosti i moguće nedostatke, a time će se i ući u razradu samog skladišta podataka koji je temelj svake poslovne analize. Kako bi se objasnilo skladište podataka potrebno je pojasniti ključne pojmove koji se odnose na rad sa skladištima podataka. Pojmovi će uključivati praktični dio rada u kojem je pohranjena već postojeća baza igrača crossmintona¹ sa stvarnim podacima, njihovim rezultatima iz sezone 2018. godine. Svi pojmovi i alati spomenuti u radu će biti potkrijepljeni primjerima iz prakse kako bi se pobliže, pomoću fotografija (engl. *Screenshot*) postojeće baze te korištenih alata objasnio smisao i proces koji se odvijao. Kako bi se prikazao konačni rezultat istraživanja potrebno je proći određeni broj koraka koji dovodi do konačne vizualizacije konkretnih podataka. Koristit će se Microsoft alati koji su danas najšire dostupni glede poslovne inteligencije, a početak će se od najosnovnijeg tipa. Tip koji je većini korisnika poznat jest Microsoft Excel koji se koristi zajedno sa svojim dodacima. Opisat će se nešto kompleksniji skup alata koje čine SQL Server. Spomenuti će se važnost ETL procesa te pojasniti na primjerima. Za konačni prikaz rezultata će biti korišten Microsoft Power BI alat koji je danas na tržištu široko uporabljiv što govori o samoj pristupačnosti i korisnosti korištenja upravo tog alata za prikaz podataka. Tim procesom se prikazuje koja kompleksnost stoji u pozadini neke vizualizacije podataka i kroz što sve se mora proći kako bi korisnik na posljetku mogao dobiti konkretne podatke na temelju kojih može provoditi daljnju analizu i provesti donošenje poslovnih odluka i time osigurati pozitivnu budućnost poslovanja.

¹ Crossminton je sport s reketom, koji igrom podsjeća na kombinaciju tenisa, skvoša i badmintona

2. Poslovna inteligencija

Poslovna inteligencija je već prilagođeni pojam u hrvatskom jeziku koji je možda više poznatiji pod kraticom BI. Poslovna inteligencija se utemeljila i kao pravilo uspješnog poslovanja, a može se objasniti kao samostalan proces prvenstveno vođen tehnologijom za analizu podataka i u konačnici uspješno prezentiranje tih podataka bilo klijentima ili za vlastito poslovanje kako bi se zaključno donijela neka odluka ili zaključak poslovanja. Poslovna inteligencija se zapravo odnosi na različite alate, aplikacije i metode koje zajedno omogućuju prikupljanje skupova podataka iz izvora koji stoje na raspolaganju ili ih je prethodno potrebno obraditi, „Poduzeća nastoje uvijek biti jedan korak ispred prikupljajući velike količine podataka, od demografije, ponašanja kupaca i vjernosti kupaca pa do financijskih i operativnih vrsta podataka. Nažalost, sami ti podaci bez organizacije i smislenog prikaza bi bili u potpunosti beskorisni kod donošenja odluka.“, (Lauer J. et al., How to Choose the Right Business Intelligence Technology to Suit Your Style). Skupove podataka je potrebno pripremiti za daljnju analizu te konačno posložiti izvještaj koji služi kao temelj za buduće poslovanje ili uvid u prošlo. Glavna uloga je upravo u povezanosti uz prikaz uspješnosti poslovanja kako bi se u konačnici mogle donijeti konkretne odluke te optimizirati poslovni procesi, a time i povećati efikasnost. Osim što se predviđaju eventualni problemi koji bi mogli nastupiti tijekom poslovanja mogu se i predvidjeti ili primijetiti koji su to aktualni marketinški trendovi kako bi se klijent mogao što uspješnije plasirati na tržište. Uz mogućnosti pogleda ili prognoze neke vrste budućnosti koja bi nastupila na temelju odluka koje se mogu tada donijeti, klijent ima uvid i u prošlo poslovanje, odnosno u povijest samih podataka koji se nalaze u skladištu podataka. Naravno gledajući skladište podataka gdje se nalaze neki prošli podaci može se uvelike pridonijeti nekom budućem poslovanju i temeljiti odluke upravo na njima. Informacije koje nam ti podaci daju na uvid su ključni pokretač uspješnog ili neuspješnog poslovanja, kao što je Covington (2016) u svom radu naveo da u današnjem tehnološko oblikovanom svijetu, nešto poput informacije može napraviti dobar posao ili u potpunosti uništiti neko poslovanje, samo ovisno o tome kako se koriste informacije.

2.1. Razvoj poslovne inteligencije

Covington (2016) u svom radu spominje kako analiza podataka nije tek novonastali pojam, već je ona prisutna oduvijek. Analiza podataka se počela koristiti još od kad je područje računovodstva predstavljeno u drevnoj Mezopotamiji u svrhu praćenja porasta, odnosno, smanjenja stada i usjeva. Iz tog se može zaključiti da se analiza oduvijek primjenjivala u nekoj mjeri. Nakon što se pojam „the Big Data“ pojavio, Roger Mougulas je 2005.g. već previdio da će to dovesti do ogromne količine podataka za čiju analizu neće biti dovoljni tada prisutni tradicionalni alati za analizu. 2010.g. Eric Schmidt je napomenuo da je količina informacija koja je nastala od samog početka civilizacije pa do 2003. godine velika otprilike 10 eksabajta² (engl. *Exabyte*), a to bi bilo jednako količini informacija koju ljudi tada, u 2010. godini, proizvedu u svega dva dana. Ta njegova izjava prikazuje koliko veliku značajku u današnjem vremenu čini informacija i koliko su bitni podaci koji se svakodnevno gomilaju, „U doba gdje informacija dominira, nije ni čudo što su upravo podaci postali najviše cijenjena valuta, a tu najviše profitiraju oni koji imaju znanje za njihovo pravilno korištenje.“, (Daniel Covington, Analytics). Covingtonova (2016) izjava dokazuje koliko su danas potrebni analitičari i sustavi koji bi kontrolirali protok i organizaciju informacija.

2.1.1. Postignute promjene

Promjene koje su nastupile nastajanjem poslovne inteligencije pod pritiskom „Big data“ su zapravo preokrenule način funkcioniranja poslovanja te dale uvid u zanemareni dio koji je već istaknut kao onaj najbitniji, jer se svako poslovanje temelji na informacijama koje odlučuju o njegovoj uspješnosti. Prvenstveno to su prednosti i velike vrijednosti koje su do te promjene dovele. Prednosti uključuju novonastalo radno mjesto, odnosno pozicija, što označava veći broj radnih mjesta, ali time i rasterećenje sektora koji su se do tada trebali baviti i ovim područjem. Način na koji se sada prikazuju podaci može se nazvati revolucionarnim, jer ti podaci su sada toliko čitljivi da jasno prikazuju željene informacije i rezultate poslovanja kojim je kasnije jednostavnije

² 1 milijun terabajta

upravljati. Visoke vrijednosti se očituju u tome da postoji mogućnost previđanja kojim se otkrivaju potencijalni problemi ili s druge strane nalaze dobre strategije za buduće i uspješnije poslovanje. Napravljeni model podataka nastaje upravo s namjerom da istraži tržište i dovede poslovanje u bolju poziciju. Covington (2016) u svom radu ističe kako analiza velike količine podataka omogućuje bilo kojem sektoru poduzeća da donese bolje i jasnije odluke. Posebice spominje važnost poboljšanja u sektorima poput zdravstvene skrbi koja danas glasi kao jedan od problematičnijih glede da svi korisnici prime odgovarajući tretman te na brz i efikasan način. Ističe kako je izrazito bitna potreba da se svim korisnicima pruži najbolje moguće rješenje na najprihvatljiviji način. Bitna odlika analize podataka se danas očituje u tome da se mogu analizirati podaci točno u vrijeme kada su oni nastali, odnosno da se to izvršava u stvarnom vremenu (engl. *In real-time*), „Kako analiza podataka napreduje svakim danim, povećava se analiza upravo onih podataka koji su nastali u stvarnom vremenu ili barem blizu stvarnog vremena. Ta sposobnost pruža izuzetno brze rezultate koji nam omogućuju donošenje odluka na temelju najnovijih podataka.“, (Daniel Covington, Analytics).

2.1.2. Mogući problemi

Kako svaka promjena donosi prednosti tako nastupaju i neki problemi. U slučaju poslovne inteligencije, problematika se javlja u smislu manjka obrazovanih ljudi koji bi na adekvatan način mogli manipulirati podacima te nedostaje pravi trening u dubokoj analizi podataka. Prema knjizi „Analytics“ iz 2016. godine, autor navodi da nedostaje između 140.000 i 190.000 ljudi koji bi mogli provoditi duboku analizu podataka, a čak oko 1.500.000 ljudi nedostaje u području analize, koji bi mogli kvalitetno analizirati podatke i pravilno interpretirati te u konačnici na temelju njih donijeti ispravnu odluku. Temeljem toga može se zaključiti da postoji i premali broj edukacija upravo u te svrhe te konačno, broj adekvatnih softvera (engl. *Software*) za takvu provedbu poslova koji se i dalje nalazi u manjini. Osim manjka radne snage, problem se očituje u onoj svakodnevnoj zabrinjavajućoj činjenici da ne postoji više osobna privatnost, te se smatra da nju upravo zanemaruju radnje koje donosi analiza podataka i svrha poslovne inteligencije. Covington (2016) spominje u svom radu da većina odvjetnika

koji brane ljudska prava navode da je zapravo analiza podataka već prerasla u neku vrstu pretjeranog zadiranja u privatnost čak i osobnih života korisnika. Tu zabrinutost se pokušava minimalizirati putem nedavno uvedenog GDPR-a (General Data Protection Regulation), prevedeno na hrvatski kao, „Opća uredba o zaštiti podataka“. GDPR je osmišljen kao novi zakon koji štiti privatnost i osobne podatke te se primjenjuje u svih 28 članica EU-a, čime se pokušava smanjiti ono, tako zvano besramno zadiranje u privatnost korisnika. Provedena je analiza od strane RSA Security-a koja je rezultirala upravo statistikom da su korisnici zaista zabrinuti za sigurnost svojih podataka te da drže do privatnosti. Ispitano je 7.500 osoba iz država Francuske, SAD-a, Italije, Njemačke i Ujedinjenog Kraljevstva te su postignuti sljedeći rezultati.

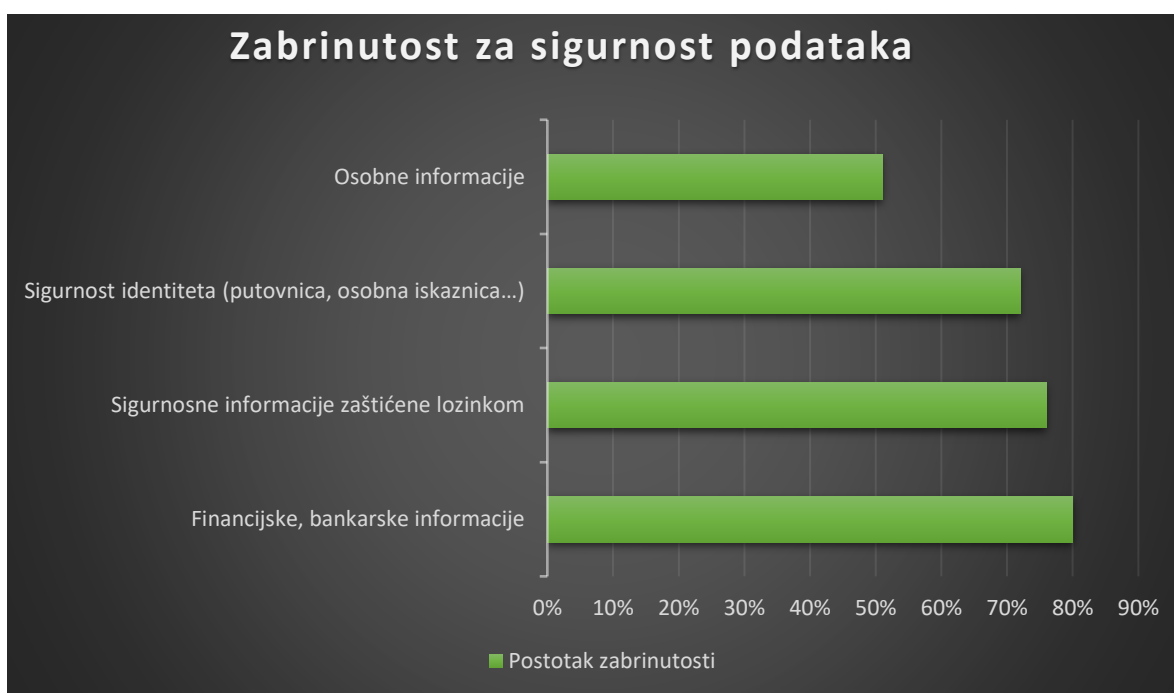


Figure 1: Postotak zabrinutosti ispitanika za pojedinu vrstu sigurnosti podataka

Graf iznad prikazuje u kolikoj mjeri su ispitanici zabrinuti za svoje podatke te vezano uz koju vrstu vlastitih podataka. Najviši stupanj zabrinutosti se odnosi na sigurnost vezanu uz podatke koje posjeduje njihova banka, dakle financije. Drugi tip je vezan uz sigurnost na internetu gdje su osobni podaci pojedinaca zaštićeni lozinkom koja može biti dosta krhka obrana od potencijalnih napadača. Svega 4% manje se odnosi na zabrinutost za sigurnost vlastitog identiteta putem putovnica, osobnih iskaznica,

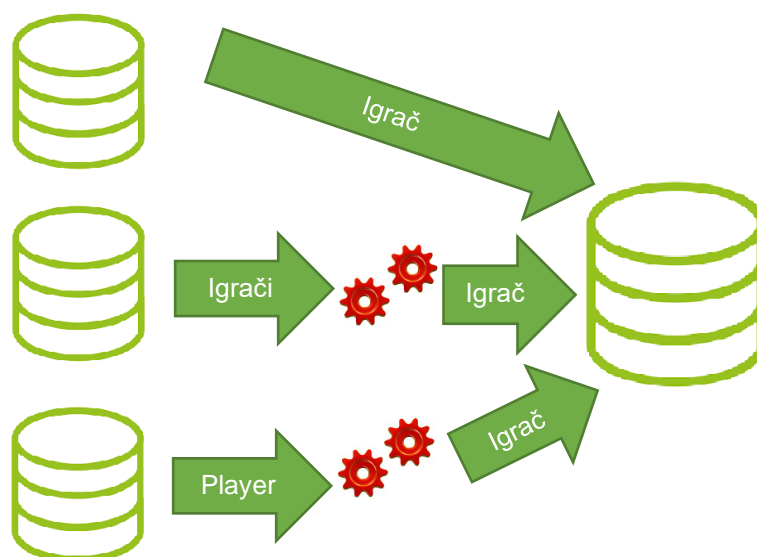
vozačkih dozvola ili nekih drugih srodnih vrsta dokumentacije. Najmanji od ponuđenih se odnosio na sigurnost vlastitih podataka koji bi služili za eventualnu ucjenu od strane hakera, za taj dio se uglavnom izjašnjavala nešto mlađa populacija u dobi od 18 do 24 godine starosti. Može se zaključiti da je ta nedoumica oko vrste podataka koji bi trebali biti obrađivani i u koje svrhe glavna briga osoba ili kompanija koji su zaduženi za to, stoga podaci koji dolaze na analizu trebaju biti pomno odabrani i validni.

3. Skladište podataka

Skladište podataka (engl. *Data Warehouse*) je zapravo temelj poslovne inteligencije koji nastoji integrirati podatke iz više različitih izvora podataka i time ubrzati izvješćivanje na način da se podaci preoblikuju kako bi korisnicima bili jednostavniji i pristupačniji. Poznatija definicija koju je opisao Bill Inmon govori da je skladište podataka skup podataka koji služi kao potpora odlučivanju, a mora biti:

- subjektivno orijentiran
- integriran
- postojan
- vremenski različit

Da je subjektivno orijentiran označava da su podaci organizirani oko subjekta kojim se bave, dakle to može biti kupac, igrač, dobavljač i slično. Radi se o podacima koji pružaju jednostavan pregled gdje se izostavljaju oni nepotrebni podaci. Integriran je onda kada su podaci prikupljeni s različitih izvora, a potom se modificiraju kako bi se pretvorili u format koji je potreban i prikladan za korištenje.



Slika 1: Primjer integracije podataka

Također mora biti postojan, podrazumijeva se da su podaci u skladištu podataka postojani, a to su ciljano podaci koji sadrže povijesne podatke te su orijentirani prema analizi podataka, za razliku od transakcijskog sustava gdje se događaju promjene.



Slika 2: Razlika između transakcijskog sustava i skladišta podataka

Na slici broj 2 je slikovito opisano koja je to razlika koju čini transakcijski sustav u odnosu na skladište podataka. Može se zaključiti da transakcijski sustav čini promjenu u sustavu, odnosno tu se događaju procesi koji uključuju punjenje i nadopunu ili brisanje podataka, gdje se koriste SQL naredbe poput „*DELETE*“, „*SELECT*“, „*INSERT*“, „*UPDATE*“, dok je skladište podataka mjesto na kojem se ne događaju promjene, već se može samo pretraživati i koristiti za daljnju analizu, odnosno koristiti SQL naredbu „*SELECT*“.

Table 1: Razlikovni odnos transakcijskog sustava i skladišta podataka

TRANSAKCIJSKI SUSTAV	SKLADIŠTE PODATAKA
Sadrži trenutne podatke	Sadrži povijesne podatke
Sadrži detaljne podatke	Sadrži detaljne i sumarne podatke
Podaci su promjenjivi	Podaci su postojani
Predvidljivi načini korištenja	Nepredvidljivi načini korištenja
Orijentiran prema dnevnim operacijama i vođenju poslovnog sustava	Orijentiran prema analizi podataka
Potpora dnevnim, operativnim odlukama	Potpora strateškim odlukama
Poslužuje velik broj operativnih korisnika	Poslužuje manji broj korisnika
Težište na pohranjivanju podataka	Težište na dobavljanju informacija

Tablica 1 prikazuje jasne razlike u smislu kako se koriste, za koga i na koji način. Upozorava na dva različita oblika koja su nužna kako bi buduća analiza bila uspješna i potpuna. Osim spomenute postojanosti, bitna je i vremenska oznaka. To je nužna odrednica jer se svi podaci analiziraju u gotovo svim situacijama upravo u vremenskom kontekstu, stoga svaki zapis ima neku vremensku oznaku. Potrebno je poznavati od kojeg do kojeg vremenskog trenutka vrijedi neki podatak, pogotovo zbog analize te spoznaje vremenskog razdoblja čuvanja podataka.

3.1. Dimenzijski model podataka

Dimenzijski model se može opisati kao proces kojim transformiramo prikaz informacija u razumljiviji i čitljiviji format. Kimball R., kao jedan od vrsnih poznavatelja upravo dimenzijskog modela podataka navodi da stvar koju prvenstveno primjećujemo kod sheme dimenzijskog modela je zapravo njegova jednostavnost i simetrija. Razlikujemo dva modela kojima se vodi modeliranje:

- Zvezdasti model (engl. *Star schema*)
- Pahuljasti model (engl. *Snowflake schema*)

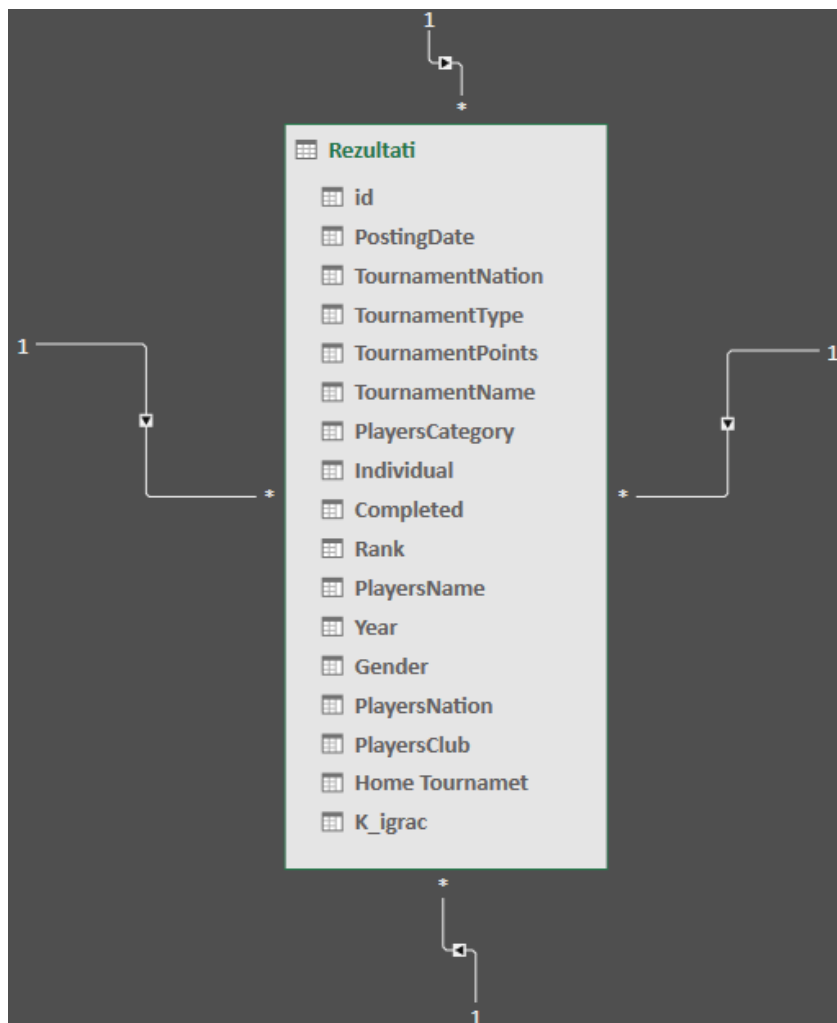
Pomoću ta dva različita modela dobije se prikaz dimenzijskog modela. Naziv je dobiven po samom izgledu shema. Upravo je najveća prednost dimenzijske strukture podataka sama mogućnost vizualizacije podataka. Uglavnom se tu radi o tri dimenzije, pa se podaci mogu shvatiti kao kocke. Zbog toga se u tom pogledu i prikazuju na taj način kako bi bile razumljivije analitičke obrade u obliku mjera (engl. *Measures*) te činjenica (engl. *Facts*). Činjenična tablica koja označava izrazito bitnu ulogu jednog dimenzijskog modela, zapravo je mjesto gdje su pohranjene numeričke vrijednosti kao što su mjere i činjenice, „...mjera i činjenica su jednostavno drugačiji nazivi za brojčane vrijednosti koje služe za analizu podataka.“ (Withee, Microsoft Business Intelligence For Dummies). Mjera naravno ne mora biti isključivo numerička vrijednost koja se može dohvatiti iz baze podataka, prema Withee (2010.), mjera bi trebala biti nešto što je prebrojivo i što se može objediniti, odnosno što bi moglo koristiti agregaciju kod sjedinjenja podataka. U odnosu na informacije koje se žele dohvatiti, korisnik može kreirati i vlastitu mjeru.

3.1.1. Kocke

Kada se povežu analitička obrada i korisničko sučelje, to se odnosi na pojam online analitičke obrade, više poznatijeg pod kraticom OLAP (engl. *Online Analytical Processing*) alata. Stvoren je kako bi se putem SQL upita moglo pristupiti bazi jer se baza periodički popunjava i treba biti konstantno u toku događanja. Kimball R. (2002.) tvrdi da bi se uloga OLAP kocki trebala pomno razmotriti, posebice dio koji služi za strategiju združivanja podataka, iz razloga što su izrazito prikladne za davanje brzih odgovora na sažete podatke. Ključna karakteristika koja čini neku kocku (engl. *Cube*) je njezina multidimenzionalnost, što označava da može određene podatke analizirati prema nekom kriteriju kao što su količina, vrijeme ili neka druga mjerna jedinica, „Kad je kocka podataka napravljena, multidimenzionalna baza podataka obrađuje oskudne podatke, dok istovremeno minimalizira one prekomjerne količine podataka te ih svodi na eksplicitne nule.“ (Kimball R., Ross M., *The Data Warehouse Toolkit*).

3.1.2. Činjenična tablica

Ključan pojam u dimenzijskom modelu podataka nosi upravo činjenična tablica (engl. *Fact table*), odnosno možda poznatiji naziv „*faktica*“, „Činjenična tablica je primarna tablica koja se nalazi u dimenzijskom modelu, a u nju se pohranjuju numeričke mjere poslovanja.“ (Kimball R., Ross M., *The Data Warehouse Toolkit*). Tablica odgovara procesu koji se prati u skladištu podataka, a samo skladište može sadržavati N broj činjeničnih tablica. Tablica se sastoji od ključeva dimenzijskih tablica i njezinih mjera.



Slika 3: Činjenična tablica

Na slici broj 3 vidi se primjer činjenične tablice koja se koristi u praktičnom dijelu rada. Prikazano je da se sastoji od nekoliko ključeva te je povezana vezom N:1 s pripadajućim dimenzijskim tablicama. Dok se ova vrsta tablica sastoji od nešto većeg broja vrijednosti, za razliku od dimenzijskih gdje ih može biti značajno manje.

3.1.3. Dimenzijska tablica

Dimenzijska tablica odgovara subjektu ili objektu koji sudjeluje u procesu koji se prati u skladištu podataka, a svako skladište može sadržavati N dimenzijskih tablica. Zapravo služi kako bi objasnila činjenice koje su pohranjene u činjeničnoj tablici. U dimenzijskoj tablici je navedena određena količina vrijednosti koja pobliže opisuje neku tablicu te pruža detaljnije informacije o podacima, „Atributi dimenzijskih tablica igraju vitalnu ulogu na mjestu skladišta podataka. Kako su dimenzijske tablice virtualni izvor svih potrebnih ograničenja i oznaka nekog izvještaja, tako one čine skladište podataka pogodnim za daljnje korištenje i razumijevanje.“ (Kimball R., Ross M., The Data Warehouse Toolkit).

2018 ICO Crossminton Robocom Swiss Open		
JAGECIC Zrinka		1
JUNKER Ivo		1
KNOBLING Charly		1
STRÖBEL Regina		1
WIRTH Frank		1
WIRTH Severin		2

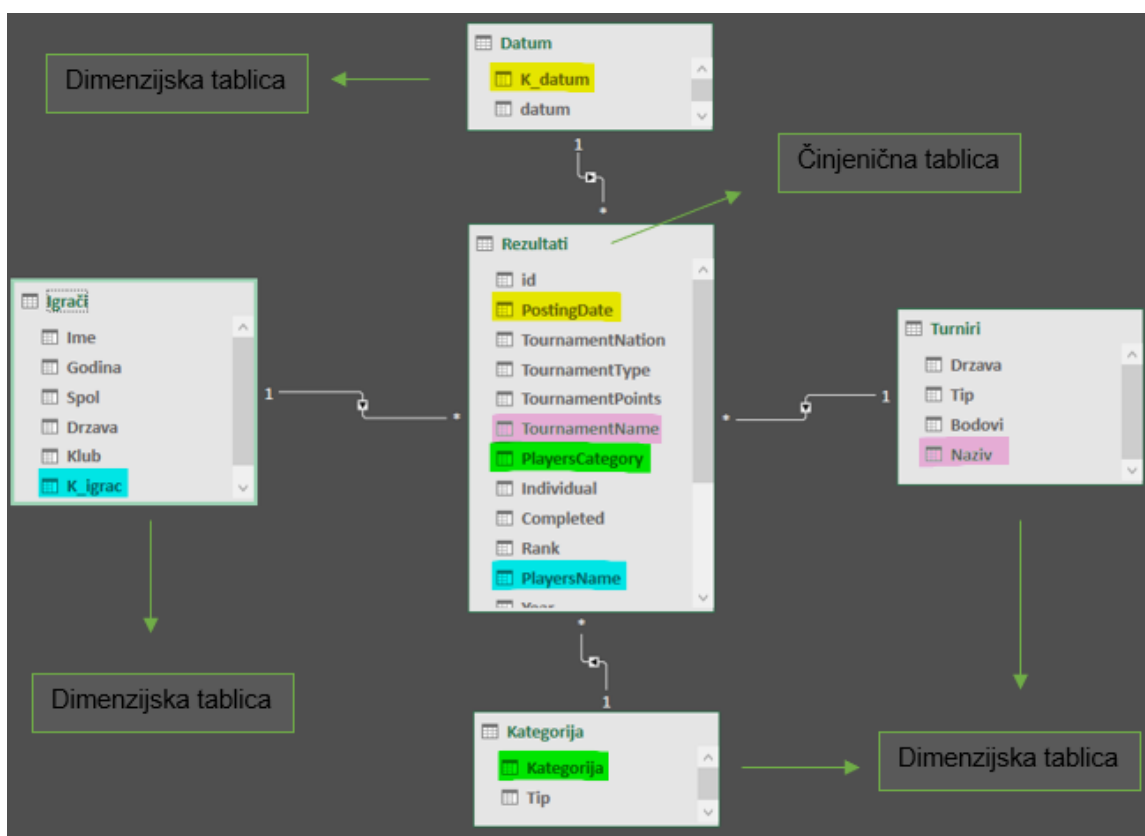
Filters		Columns	
Rank			
Rows		Σ Values	
TournamentName		Count of PlayersName	
Ime			

Slika 4: Funkcija dimenzijske i činjenične tablice

Princip po kojem dimenzijska tablica opisuje radnju činjenične tablice se može iščitati sa slike broj 4. Na slici su prikazana imena igrača i igračica koji su osvojili prva mjesta na nekom turniru, u ovom slučaju „2018 ICO Crossminton Robocom Swiss Open“, „Ime“ dolazi iz dimenzije „Igrač“, a „Rank“ iz dimenzije „Kategorija“, također u obzir je uzeta i vrijednost gdje se vrši izračun „PlayersName“, odnosno, koliko je puta ime igrača upisano pod „Rank 1“ kako bi se dobio broj prvih mjesta pojedinog igrača. Stoga, da bi se došlo do željenog rezultata gdje se želi doznati koji igrač je osvojio prvo mjesto na tom turniru, potrebno je iz činjenične tablice pod nazivom „Rezultati“ pogledati polje „PlayersName“ koje će konačno dati rezultat turnira. Shodno tome, može se zaključiti da je činjenična tablica upravo ovisna o dimenzijskoj kako bi se shvatio njen kontekst.

3.1.4. Zvezdasti model

Zvezdasti model (engl. *Star schema*) u dimenzijskom modeliranju skladišta podataka je najčešće korišten oblik, a također predstavlja i najjednostavniji tip koji svojim izgledom podsjeća na zvijezdu. U središtu samog modela je smještena činjenična tablica, ona je vezom N:1 povezana s dimenzijskim tablicama koje se nalaze u njenoj okolini.



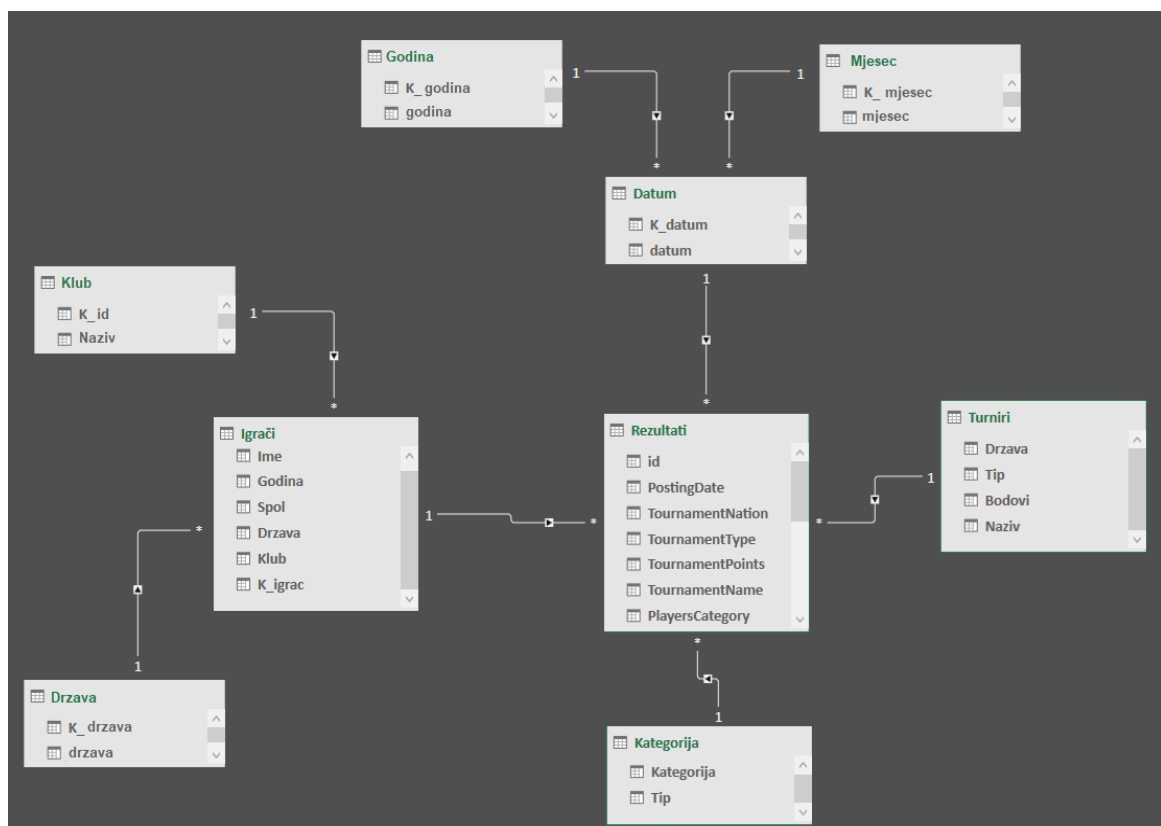
Slika 5: Primjer zvezdastog modela (engl. *Star schema*)

Na slici 5 je prikazano kako to u praksi izgleda kada se radi veza između tablica u modelu skladišta podataka. Različitim bojama su označeni atributi u činjeničnoj tablici koji su zapravo ključevi pripadajućih dimenzijskih tablica. Dimenzijske tablice koriste strani ključ (engl. *Foreign key*) kako bi se povezale na činjeničnu tablicu, no valja primijetiti da dimenzijske tablice nisu međusobno povezane, upravo to odražava koliko

je bitna uloga činjenične tablice. Ova shema je podržana u svim alatima za poslovnu inteligenciju i jednostavna je za korištenje.

3.1.5. Pahuljasti model

Pahuljasta shema (engl. *Snowflake schema*) je zapravo produbljena verzija zvjezdaste sheme koja ima dodatne dimenzije. Ime je dobila upravo po obliku na koji podsjeća. Slika 6 je primjer pahuljastog modela gdje mogu se primijetiti dodatne 4 dimenzije. Dimenzije su podijeljene tako da se dvije vežu na dimenziju "Datum", a dvije na dimenziju "Igrači". Razlika je što ovdje svaka dimenzija može imati svoju posebnu dimenziju. Vidljivo je da su podaci otišli još više u dubinu razrade, a time predstavljaju nešto kompleksniji dizajn baze podataka. Naravno povećanjem broja dimenzija povećava se i broj pridruživanja tablica, odnosno joinova kako bi se došlo do konačnog rezultata kod procesa dohvaćanja podataka.



Slika 6: Pahuljasta shema (engl. *Snowflake schema*)

3.1.6. SQL

SQL je kratica za „Structured Query Language“, kao što i samo ime govori, vrsta računalnog jezika za strukturiranje podataka. Ta vrsta jezika se može definirati kao računalni jezik koji se koristi za pohranjivanje, manipuliranje i preuzimanje onih podataka koji se nalaze pohranjeni u nekoj relacijskoj bazi podataka. Zapravo je SQL standardni računalni jezik stvoren za relacijske baze podataka. Također glasi kao standardni jezik za sve baze podataka poput:

- MySQL Server
- Microsoft Access
- Oracle
- Sybase
- Informix
- SQL Server

Koristi se zbog svojih pogodnosti za korisnika. Prvenstveno se odnosi na to da dozvoljava korisnicima pristup podacima u RDBMS-u (engl. *Relational Database Management System*), prevedeno na hrvatski, sustav za upravljanje relacijskom bazom podataka. Korisnici mogu definirati podatke u bazi podataka i manipulirati njima, no ono najbitnije je da može kreirati i uklanjati baze podataka i tablice unutar baze, a to su najvažnije karakteristike koje označuju jedan SQL programski jezik za poslovnu inteligenciju. Podaci su, u RDBMS-u, kao i kod uobičajenog načina pohranjivanja podataka, spremljeni u tablice koje su raspoređene u kolone i redove.

	Ime	Godina	Spol	Drzava	Klub
1	YAMAGUCHI Yuri	1903	F	JPN	Shinagawa Club
2	VASKO Stefan	1945	M	SVK	Lokomotiva Banska B.
3	SCHÜSSELER Manfred	1947	M	GER	Brühler TV-Speedkrebse Köln
4	KOZIELSKI Czeslaw	1950	M	POL	UKS MILESZA Targowa Gorka
5	ROSSI Frank	1951	M	SUI	Speeders Club Ticino
6	VIKRAM DATT Shama	1951	M	IND	Phoenix Sports Club
7	WINCENCIAK Jan	1951	M	POL	KC Wolf Fighters
8	BLANUŠA Vladimir	1952	M	SRB	Speeder Sombor
9	IVANOVIC Tomislav	1952	M	SRB	Speeder Sombor
10	MONFILS Martin	1953	M	NED	RAP Speeders

Slika 7: Tablica u SQL-u

Slika broj 7 prikazuje kako izgleda neka tablica u SQL-u te što predstavljaju redovi, a što kolone tablice. Prikazuje kako se tablica sastoji od manjih entiteta koji su smješteni u polja, a polja su organizirana tako da predstavljaju neku specifičnu informaciju koja je sadržana u toj koloni. Podaci se mogu dobiti postavljanjem konkretnih naredbi, a te naredbe se mogu podijeliti u 3 grupe, ovisno o tome radi li se o definiranju podataka, manipuliranjem već postojećih podataka ili o tome tko ima i koliku kontrolu nad njima. Kako bi bilo jasnije zašto je potrebno poznavati razliku između te tri vrste naredbi, prikazat će se i konkretni primjeri kod jednostavnije uporabe:

- DDL (Data Definition Language)

- CREATE – kreira novu tablicu, pogled (engl. *View*) ili objekt u bazi podataka

```
create table NovaTablica (  
  id int NOT NULL,  
  Ime nvarchar (50) NOT NULL,  
  Ranking int,  
)
```

- ALTER – modificira već postojeći objekt unutar baze podataka, kao što je tablica

```
alter table NovaTablica add  
Kategorija varchar NOT NULL
```

- DROP – briše čitavu tablicu ili pogled unutar baze podataka

```
drop table NovaTablica
```

- DML (Data Manipulation Language)

- SELECT – preuzima i prikazuje tražene podatke iz jedne ili više međusobno povezanih tablica

```
select * from NovaTablica
```

- INSERT – ubacuje nove ili već postojeće zapise

```
insert into NovaTablica (id, Ime)  
values (1, 'Zrinka')
```

- UPDATE – izmjenjuje podatke

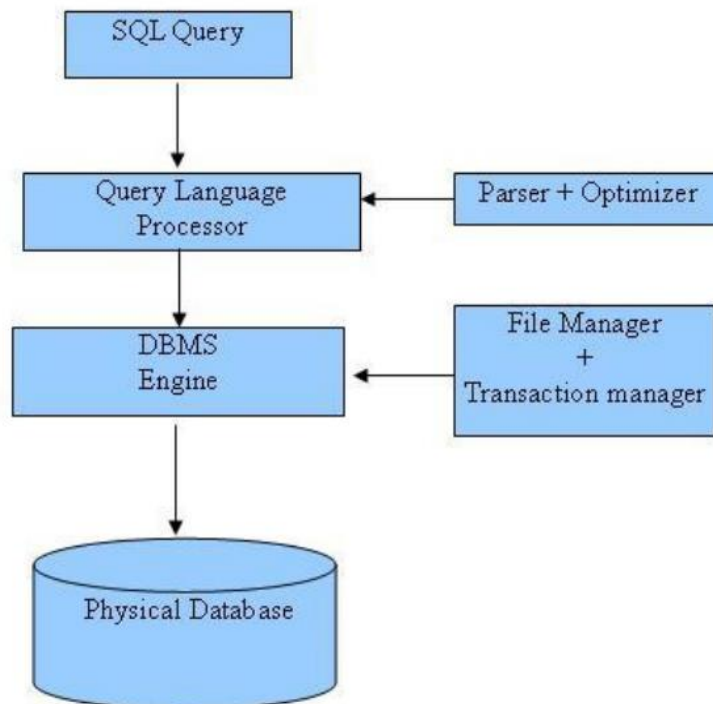
```
update NovaTablica  
set Ime='Laura'
```

- DELETE – briše zapise

`delete from` NovaTablica

- DCL (Data Control Language)
 - GRANT – pridaje privilegije nekom korisniku
 - REVOKE – uklanja prethodno pridate privilegije

Slika broj 8 daje uvid u to da je SQL izrazito inteligentan proces koji prije nego što izvrši naredbu prolazi kroz optimizaciju upita te na temelju toga pronalazi najbolji mogući način kako bi mogao proslijediti naredbu do fizičke baze podataka te prikazati konačni rezultat.



Slika 8: SQL arhitektura

4. Microsoft alati

Microsoft ima alate koji uvelike pomažu analitičarima, menadžerima ili developerima kako bi svoje radne zadatke i ciljeve izvršili što brže, efikasnije i sa što većom preciznošću. Bitno je poznavati koji alat je prikladan za koji aspekt izvođenja poslova te na koji način ga koristiti kako bi se postigao što bolji rezultat za poslovanje. Microsoft pruža široki spektar alata koji su na raznim težinskim razinama što se tiče manevriranja njima. Tako je široko poznat i relativno jednostavan za korištenje, MS Excel, kao značajan dio Microsoft Officea, a svi se korisnici kad tad susreću s čitavim paketom u kojem je upravo MS Excel izrazito često korišten, kao nezaobilazan alat svakog poslovanja. Drugi izrazito poznati alat koji iziskuje malo veći spektar poznavanja programskih jezika, odnosi se na Microsoft SQL Server koji je vodeći u području stvaranja i održavanja baza podataka. Konačno, alat namijenjen za vizualizaciju podataka je sve više poznati Power BI. U daljnjem radu će se prikazati svi ovi spomenuti alati, te će se pružiti i praktični primjer na konkretnim podacima.

4.1. Microsoft Excel

MS Excel je izrazito široko rasprostranjen alat, posebice u vidu „samoposlužnog“ (engl. *SelfService*) načina analize. Relativno jednostavan za korištenje pošto je svaki korisnik imao barem jednom priliku upoznati se s ovim alatom. Omogućava brzu izradu analize i izvještaja na nešto jednostavniji način za razliku od drugih alata. Sam po sebi je dosta fleksibilan alat koji omogućava izradu raznih tipova analize koje bi se mogle izvršiti, a uz to daje izrazitu slobodu korisnicima da izorganiziraju podatke na način na koji žele te uključe razne izračune kako bi pružili prikaz podataka na brži način bez većih komplikacija u sastavljanju. Izrazito jednostavno se može podijeliti izvještaj s klijentima ili suradnicima, također omogućava zamrzavanje kolona kako ne bi došlo do neželjenih promjena od strane klijenata ili pogrešnih manipulacija, stoga je postignuta dobra razina sigurnosti. Alat koji je dizajniran u smislu da organizira, analizira i vizualizira podatke. Većina stručnjaka ga zapravo opisuje kao jedan od moćnijih alata na tržištu, „Excel je jedna od najmoćniji desktop aplikacija koje se

nalaze u arsenalu Microsoft BI-a“ (Withee, Microsoft Business Intelligence for Dummies).

1	PlayersCategory	Individual	Rank	PlayersName	Year	Gender	PlayersNation
2	Crossminton female O50 Individual	D	7	BOMIRSKA Lidia	1957	F	POL
3	Crossminton female O50 Individual	D	7	KAPLON Joanna	1964	F	POL
4	Crossminton female O50 Individual	D	10	MORAWSKA Anna	1957	F	POL
5	Crossminton female O50 Individual	D	10	SEK Krystyna	1958	F	POL
6	Crossminton female O40 Individual	D	5	JAGIELLO Beata	1974	F	POL
7	Crossminton female O40 Individual	D	10	JURKIEWICZ KOSIŃSKA Sylwia	1978	F	POL
8	Crossminton female Adults Individual	D	3	WIKLANSKA Natalia	1998	F	POL
9	Crossminton female Adults Individual	D	5	SCAVNICKA Maria	1976	F	SVK
10	Crossminton female Adults Individual	D	7	KAPLON Joanna	1964	F	POL
11	Crossminton female Adults Individual	D	10	WOJCIK Monika	1959	F	POL
12	Crossminton female Adults Individual	D	10	MORAWSKA Anna	1957	F	POL
13	Crossminton male O40 Individual	D	1	CHMIELEWSKI Grzegorz	1970	M	POL
14	Crossminton male O40 Individual	D	3	CWIAKALA Dariusz	1970	M	POL
15	Crossminton male O40 Individual	D	5	SAWICKI Mariusz	1977	M	POL
16	Crossminton male O40 Individual	D	7	KOWALCZYK Mariusz	1977	M	POL
17	Crossminton male O40 Individual	D	10	JAGIELLO Wojciech	1975	M	POL
18	Crossminton male O40 Individual	D	10	WINCENIAK Jan	1951	M	POL
19	Crossminton male O40 Individual	D	10	JANČUNSKI Andrzej	1978	M	LTU

Slika 9: MS Excel primjer jednostavne baze igrača

Slika broj 9 prikazuje primjer jednostavnije baze igrača. Kao što se može iščitati vidljivo je koju kategoriju igrač igra, njegovo ime i prezime, godinu rođenja, spol te nacionalnost. Osim ovih već karakterističnih podataka koje bi baza trebala sadržavati nalaze se još i kolone koje svrstavaju igrače po vrsti nastupa igrača, da li individualni nastup ili parski, stoga je tu odrednica pod „D“ i „N“. Osim vrste, naravno mora biti vidljiva i pozicija igrača koja je ostvarena na nekom turniru.

TRANSPOSE				=IF(LEFT([@PlayersNation],3)=[@TournamentNation], "D", "N")			
	C	K	N	P			
1	Tour	PlayersName	Playe	Home Tournament			
2	POL	BOMIRSKA Lidia	POL	=IF(LEFT([@PlayersNation],3)=[@TournamentNation], "D", "N")			
3	POL	KAPLON Joanna	POL	D			
4	POL	MORAWSKA Anna	POL	D			
19	POL	JANČUNSKI Andrzej	LTU	N			
20	POL	JANČUNSKI Kamil	LTU	N			

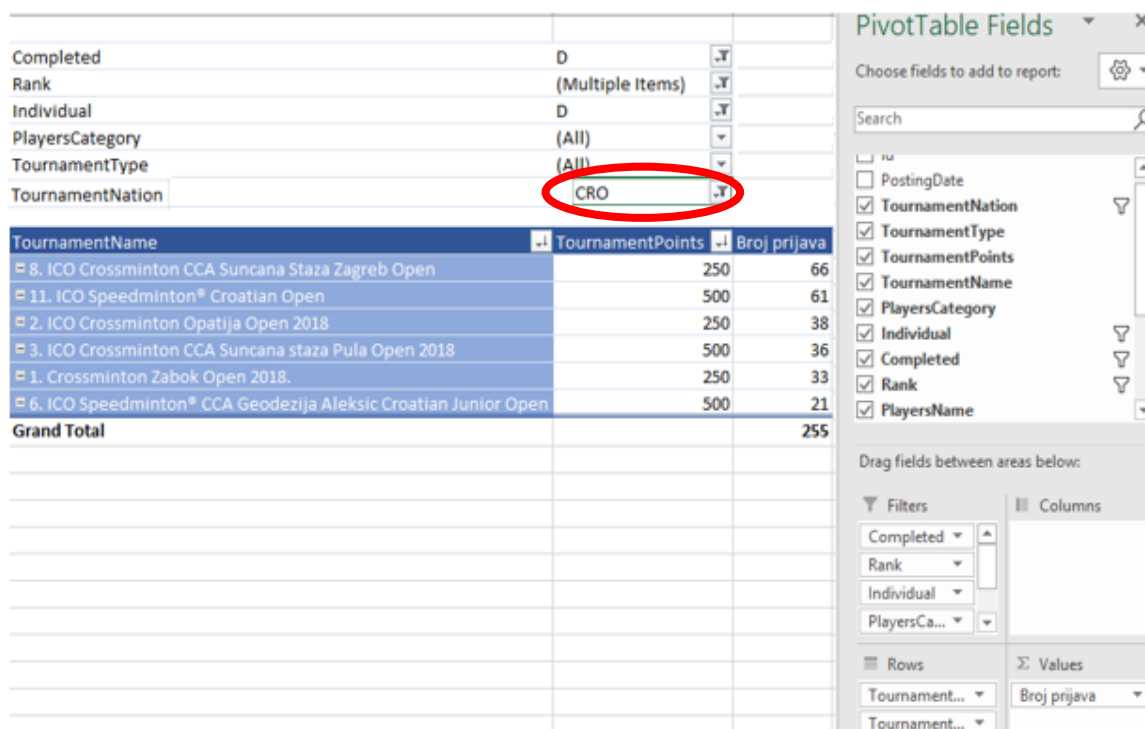
Slika 10: MS Excel prikaz mogućnosti automatizacije zapisa

Slika broj 10. daje uvid u mogućnosti automatizacije u nekom izvještaju pomoću MS Excela. Može se primijetiti da se kolona pod nazivom „Home Tournament“ može automatizirati tako da se provjerava ako se neka kolona podudara ili ne. U tom slučaju

se gleda kolona „Tournament Nation“ s kolonom „Player Nation“ i temeljem toga na znak poklapanja ispisuje oznaku „D“ za „Home Tournament“ ili u suprotnom „N“, kada to nije slučaj. Na ovaj način popunjavanje baze je itekako ubrzano i pojednostavljeno, nije potrebno sirovo upisivanje podataka, već MS Excel sam prepoznaje kriterije i sam ispisuje zadanu oznaku što je ujedno i cilj, kako bi se uredsko poslovanje ubrzalo i automatiziralo.

4.1.1. Pivot tablica

Sami MS Excel je podosta jednostavan, no iziskuje formule kako bi se moglo doći do prikaza podataka i postavljanja konačnog izvještaja. Kao velika pomoć tu je osmišljena takozvana Pivot tablica kao izrazito inteligentno rješenje koje radi po jednostavnijem principu. Tablica može razdvajati dijelove i sortirati ih kako joj se dodijeli odrednica, bez formula ili znanja SQL-a. Bitno je jedino poznavati logiku po kojoj funkcionira. Pivot tablica omogućuje da se reorganizira zapis podataka u njoj i konačno sumira željeni iznos prema bilo kojim kriterijima.



TournamentName	TournamentPoints	Broj prijava
8. ICO Crossminton CCA Suncana Staza Zagreb Open	250	66
11. ICO Speedminton* Croatian Open	500	61
2. ICO Crossminton Opatija Open 2018	250	38
3. ICO Crossminton CCA Suncana staza Pula Open 2018	500	36
1. Crossminton Zabok Open 2018.	250	33
6. ICO Speedminton* CCA Geodezija Aleksic Croatian Junior Open	500	21
Grand Total		255

Slika 11: MS Excel Pivot tablica

Na slici broj 11. vidljiv je način po kojem funkcionira Pivot tablica te zašto je toliko bolje rješenje za izradu izvještaja. Vidljivo je sortiranje turnira gdje je poljem „TournamentNation“ ograničeno isključivo na turnire na području Hrvatske, „CRO“. Zatim ono što je bila želja dobiti jest, broj pojedinačnih prijava na sve turnire u Hrvatskoj u sezoni 2018. S desne strane Excela nalazi se Pivot polje koja je potrebno organizirati na željeni način. Kao vrijednost, „Values“ nalazi se „Broj prijava“ koji su određeni redovima „TournamentName“ i „TournamentPoints“, jer je želja dobiti broj prijava po turniru. Za filtriranje podataka, kako bi se dobila željena i točna informacija, uzeto je u obzir da bude „Completed“ s oznakom „D“ te da se uzimaju u obzir isključivo pojedinačne prijave, ne uključujući parske, dakle, u obzir dolazi polje „Individual“ samo s oznakom „D“.

4.1.2. Power Pivot

Uzevši u obzir što sve MS Excel pruža, vidljivo je da se on za nešto manju količinu podataka postavlja kao izrazito dobro i jednostavno rješenje kod izrade izvještaja i postavljanja vizualizacije. Valja uzeti u obzir da sami Excel ima limit od svega milijun redova po stranici (engl. *Sheet*), stoga nije prikladan za korisnika koji barata s ogromnim količinama podataka. MS Excel je alat prikladan za krajnjeg korisnika koji može dalje manipulirati podacima, izrađivati izvještaje, pretvarati podatke u vizualne figure kako bi bile čitljivije, ali i povezati se na skladišta podataka, kocke podataka i ostale vanjske izvore podataka koji bi eventualno bili potrebni korisniku za bolje postavljanje konačnog izvještaja i uspješnijeg poslovanja. Zbog tog smanjenog kapaciteta Microsoft je uz već spomenutu Pivot tablicu unutar samog MS Excela, 2010. je osmislio dodatni alat na Excel pod nazivom Power Pivot. Ima veći kapacitet od samog Excela, no i dalje je limitiran na broj od oko 2GB pohrane. Bez obzira na to, Power Pivot je odlično rješenje za malo skladište podataka na kojem će se temeljiti analiza poslovne inteligencije. Power Pivot se može opisati kao alat namijenjen za stvaranje podatkovnih modela, uspostavljanje odnosa i stvaranje izračuna, ono omogućuje stvaranje većih skupova podataka koji su složenih odnosa i konačno složene ili jednostavne izračune. Podaci se mogu povući iz raznih izvora, također

omogućava da se osvježi (engl. *Refresh*) stanje koje se nalazi u nekoj vanjskoj bazi kako bi bilo sigurno da su izmijenjeni podaci u konačnici točni.

The screenshot shows a Power Pivot table with columns: Ime, Godina, Spol, Drzava, Klub, and K_igrac. The formula bar at the top displays the DAX formula: `=Igraci[Ime]&"-"&Igraci[Godina]`. The table contains three rows of data, with the 'K_igrac' column showing concatenated values: 'MAKRLÍK Petr-1965', 'MAKRLIK Petr-1993', and 'MAKRLIKOVA Dana-1967'. The first two rows are highlighted with green circles.

	Ime	Godina	Spol	Drzava	Klub	K_igrac
543	MAKRLÍK Petr	1965	M	CZE	Výcho...	MAKRLÍK Petr-1965
544	MAKRLÍK Petr	1993	M	CZE	Výcho...	MAKRLIK Petr-1993
545	MAKRLÍKOVÁ Dana	1967	F	CZE	Výcho...	MAKRLIKOVA Dana-1967

Slika 12: Power Pivot

Na slici broj 12 je prikazan primjer kako izgleda Power Pivot u praksi te kako se može dobiti precizniji izračun kolone, koristeći jezik DAX. DAX je jezik Power Pivota koji se koristi za pisanje formula te je namijenjen relacijskim podacima i analizi konteksta koji su dostupni. U primjeru na slici broj 12 dodana je kolona „K_igrac“ kojom će se spriječiti da se dobije krivi izračun na kraju ako postoje dvije ili više osoba s istim imenom i prezimenom, kao u ovom slučaju s „MAKRLÍK Petr“. Stoga je konkatencijom kolone „Ime“ i kolone „Godina“ uklonjena mogućnost da se prebroje dva različita igrača kao jedan te isti.

4.2. Microsoft SQL Server

SQL Server je relacijska baza podataka za upravljanje sustavima (engl. *Relational database management system*), više poznatiji pod kraticom RDBMS. Proizvod je namijenjen za shvaćanje problema koji se redovito javljaju u svakodnevnom poslovanju modernog doba, osim toga glavna zadaća mu je rješavanje upravo tih problema. Alat je od početne funkcije upravljanja tradicionalnim bazama podataka dosegno razinu gdje može osim toga obavljati puno širi spektar poslova. Withee (2010.) navodi u svom radu kako se SQL Server razvio i proširio toliko da sad već

uključuje neke dodatne značajke i funkcionalnosti koje su sposobne izdvojiti, manipulirati, pohraniti, analizirati te izvještavati o podacima svoje korisnike. Glavni zadaci koje SQL Server izvršava uključuju:

- izradu baze podataka
- upravljanje bazom podataka
- analizu podataka kroz SQL Server Analysis Services (SSAS)
- generiranje izvještaja kroz SQL Server Reporting Services (SSRS)
- izvođenje ETL proces kroz SQL Server Integration Services (SSIS)

SQL Server postoji u nekoliko izdanja. Svako izdanje je određeno za neku konkretnu primjenu, tako postoji i izdanje koje odgovara upravo temi ovog rada, poslovna inteligencija, „*Business Intelligence Edition* je novi dio koji je predstavljen u SQL Serveru 2012. godine. Ovo izdanje uključuje sve sposobnosti koje je server imao u *Standard* izdanju, no sada je prošireno kako bi zadovoljio i korisnike poslovne inteligencije te sada ima mogućnost korištenja Power View i Power Pivot dodataka.“, (MS SQL Server, *tutorialspoint*, 2016).

Alat koji SQL Server administratori koriste za upravljanje bazom podataka se krije pod kraticom SSMS (engl. *SQL Server Management Studio*), „SQL Server Management Studio je jedna komponenta radne stanice ili klijent alata koji će biti instaliran ukoliko se izabere ta komponenta u koracima instalacije. Ta sposobnost omogućuje povezivanje i upravljanje SQL serverom iz grafičkog sučelja umjesto da se mora koristiti naredbena linija.“ (MS SQL Server, *tutorialspoint*, 2016). SSMS je upravo taj faktor koji je zaslužan za provođenje zamisli nekog administratora. Withee (2010.) u svom radu navodi da SQL Server obuhvaća razne aspekte zadataka koji bi se mogli pronaći u poslovima upravljanja podataka. Autor navodi mogućnosti poput konfiguracije, opće upravljanje bazama podataka, administracija, sigurnost te konačno i dohvaćanje podataka, a time se može pretpostaviti da ovaj alat uvelike olakšava posao developerima. Jedna od prepoznatljivijih i značajnijih sposobnosti koje SSMS posjeduje je takozvani *IntelliSense*. *IntelliSense* je velika pomoć korisnicima kako bi bili sigurni da je SQL sintaksa ispravna. Ova sposobnost programa omogućuje ispisivanja, odnosno, nudi koje su to moguće sintakse na samo prvo upisano slovo. Withee (2010.) upućuje na jedan od glavnih razloga zašto je *IntelliSense* izrazito korisna značajka za administratore baza podataka. Autor ističe činjenicu da

administratori svakodnevno daju preko nekoliko stotina imena objektima u svojim bazama, stoga ih je i izrazito teško zapamtiti, a IntelliSense tu dolazi kao velika pomoć za lakše snalaženje i dohvaćanje pravog imena. Osim što olakšava pronalaženje vlastitih imena, također označava posebnom bojom one riječi koje su već ugrađene kao dio SQL jezika. „*SELECT*“ je jedna od najpoznatijih i takvi tipovi riječi se nazivaju ključne riječi (engl. *Keywords*) programskog jezika. Ključne riječi su označene drugačijom bojom kako korisnik ne bi upotrijebio u pogrešnim situacijama i time uzrokovao pogrešku (engl. *Error*).

4.2.1. SQL Server Database Engine

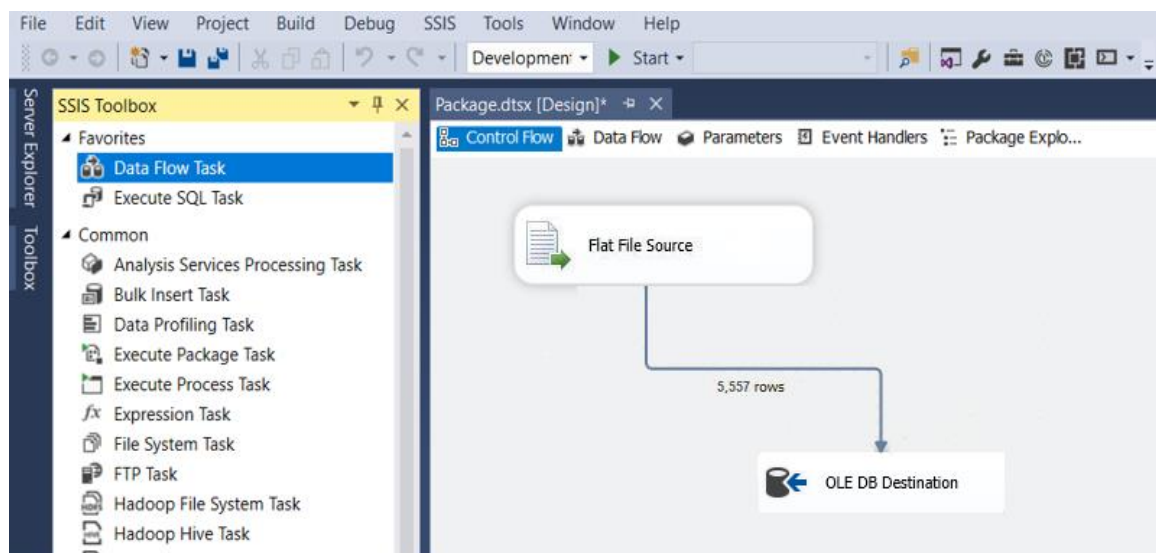
Ovaj dio SQL-a je vjerojatno najpoznatija komponenta koju ima SQL server, a i moglo bi se reći i najbitnija, ona stvara bazu podataka. Withee (2010.) u svom radu ističe da je vjerojatno ova velika uloga SQL Server Database Enginea koju on nosi, upravo razlog što ljudi često prvotno pomisle na ovaj dio SQL-a kada se spomene SQL Server.

4.2.2. SQL Server Reporting Services

SSRS je zapravo dodatak na SQL Server iz 2004. Ovaj alat služi za stvaranje izvještaja, no, danas se i ne koristi u tolikoj mjeri, jer postoje alati koji su jednostavniji i napredniji za takvu vrstu posla. Zapravo se uz SSRS za konačno izrađivanje izvještaja morao koristiti dodatak, *Report Builder*. *Report Builder* je aplikacija koja omogućuje krajnjim korisnicima jednostavno i samostalno izrađivanje izvještaja. Aplikacija je dizajnirana baš za krajnje korisnike, a kako bi se došlo do tog dijela gdje treba samo izraditi izvještaj prvo je potrebno provoditi dodatne implementacije.

4.2.3. SQL Server Integration Services

Prije nego li se podaci mogu pretvoriti u informaciju koju će netko koristiti ona se mora nalaziti u takvom formatu da bude prilagođena za pojedinu vrstu korisnika. Upravo iz tog razloga bitno je provoditi ETL (engl. *Extract, Transform, Load*) sustav koje bi prevedeno označavao ekstrakciju, transformaciju i učitavanje podataka. ETL je sustav temeljnog skladištenja podataka, stoga je potrebno imati izrazito kvalitetni sustav kako bi rezultirao ispravnim i kvalitetnim podacima. ETL se može poistovjetiti sa strategijom međujezika ³(lat. *Interlingua*), zato što je u konačnici cilj dobiti sjedinjeni set podataka, „U poslovnoj inteligenciji jezik koji bi mogao govoriti više jezika se upravo naziva ETL alat.“, (Withee, Microsoft Business Intelligence For Dummies). Sami ETL nije vidljiv krajnjim korisnicima, no on je zaslužan za onaj ključni dio izrade skladišta podataka. Microsoft je stoga razvio SSIS, kao alat koji se koristi upravo za već spomenuti ETL proces, kada je potrebno povezati se na neki izvor podataka i preuzeti te podatke na vlastiti server gdje će se oni u budućnosti obrađivati. SSIS omogućuje da se podaci preuzimaju iz raznih izvora, bez obzira na oblik u kojem se nalazili, bilo da je to bio neki Excel, baza podataka, obična tekstualna datoteka ili nešto drugo. Ovaj alat ima tu mogućnost da im pristupi, preuzme podatke i nakon toga kreće proces transformacije podataka. U konačnici kreće postupak učitavanja podataka u skladište podataka s prilagođenim podacima.



Slika 13: SSIS (engl. *SQL Server Integration Server*)

³ Strategija prijevoda, međujezik je umjetni jezik koji je neovisan o prirodnom jeziku ili njegovoj sintaktičkoj strukturi, a koristi se između prijevodnog i ciljanog jezika (prevodi s izvornog u međujezik te iz međujezika u ciljani jezik).

Slika pod brojem 13 prikazuje kako se u praktičnom dijelu rada MS Excel dokument s podacima učitao u skladište podataka gdje će kasnije biti moguće pristupiti podacima i manipulirati njima. Podaci su pohranjeni u izrađenu bazu podataka preko koje će se odvijati daljnji procesi.

4.2.4. SQL Server Analysis Services

SSAS je također jedna od komponenti SQL Servera koji se koristi kako bi se analizirali podaci. To je analitički alat koji se koristi kao potpora poslovanju i samoj poslovnoj analizi. Radi se o alatu koji je srodan već spomenutim alatima kao što je MS Excel, a to podrazumijeva i druge alate za analizu te posebno jedan alat koji je postao izrazito popularan posljednjih nekoliko godina, Power BI.

4.3. Power BI

Kako je Microsoft 2009 lansirao poznati Power Pivot kao rješenje za poslovnu inteligenciju, korisnici su željeli nešto više, odnosno, želja se javila kod krajnjih korisnika koji su imali potrebu za pristupačnijim i zapravo neovisnim alatom. Kako je Power Pivot nastao u sklopu MS Excela 2010, tako je bio ovisan o njemu, a ako se želi podijeliti vizualizacija to je ovisilo o korištenju SharePointa. Problem se javio kada je sve više korisnika prionulo današnjem modernom, mobilnom korištenju, što tada nije bilo moguće, stoga su svi ti noviteti iziskivali pojavu novog alata, a to bi bio Power BI. Tim Microsofta je osmislio ovaj alat upravo slušajući i prikupljajući ideje i zahtjeve svojih korisnika, „Korisnici su željeli mobilno iskustvo, stoga ga je razvojni tim stvorio. Analitičari podataka su željeli moć, jednostavnost i nove vizualizacije, zato je sada to sve dostupno u jednom alatu, novom alatu poslovne inteligencije, Power BI.“ (Ferrari, Russo, Introducing Microsoft Power BI). Također kreatori su se osvrnuli na tu činjenicu gdje se zapravo danas pohranjuje najveća količina podataka, a to je, u oblaku (engl. *Cloud*), stoga su odlučili stvoriti nešto što će podupirati, za razliku do sada, i takav način preuzimanja i dijeljenja podataka, “Puno truda je uloženo u stvaranje neovisnog rješenja za korisnike kako bi mogli učitati podatke s različitih izvora, posebice iz

oblaka, te zbog toga stvoriti jednu infrastrukturu koja će svim BI korisnicima omogućiti bolje izvještavanje i dijeljenje izvještaja s cijelim timom i to na način da bude jednostavno, a opet djelotvorno.“ (Ferrari, Russo, Introducing Microsoft Power BI). Microsoft Power BI je nastupio na tržište 2015 godine, a mogu se izabrati dva načina njegova djelovanja:

- Power BI Desktop za Windows
- Power BI

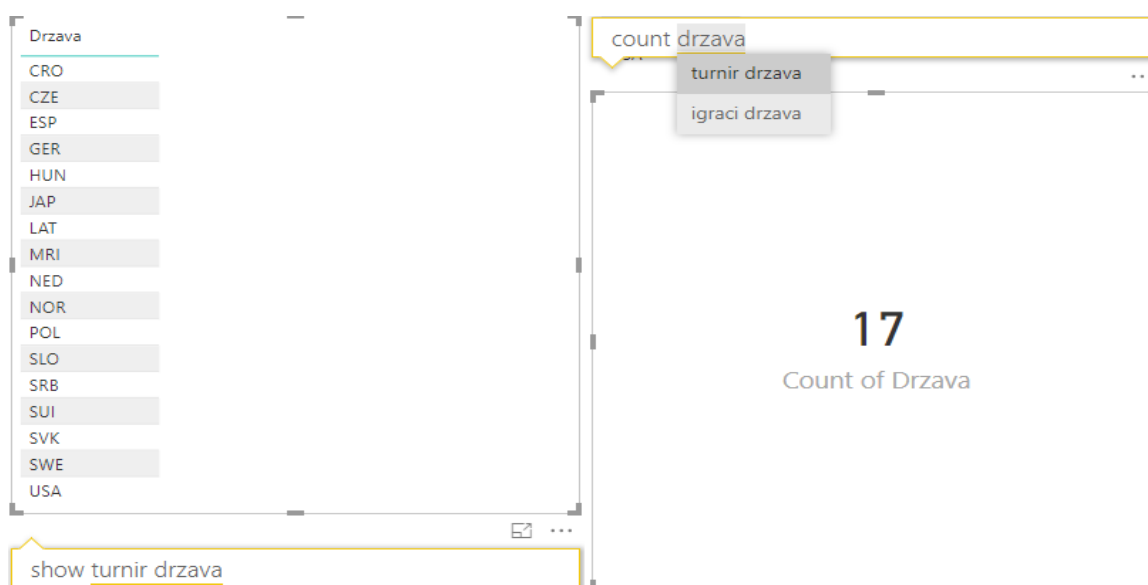
Zapravo se radi o dva gotovo identična načina rada, jedino je ona bitna razlika u tome što se aplikacija Power BI Desktop za Windows odnosi isključivo za korištenje na osobnom računalu, a Power BI je usluga namijenjena za korištenje u oblaku te se koristi preko nekog web poslužitelja. Neovisno o vrsti aplikacije, oba načina nude jednake operacije, stoga i u većini slučajeva se primjenjuju i jedan i drugi način, „Ova dva načina rada zapravo nadopunjuju jedan drugog, stoga je vrlo vjerojatno da će korisnici koristiti oba za stvaranje izvještaja.“ (Ferrari, Russo, Introducing Microsoft Power BI). Kako je Power BI web poslužitelj, za pristup je potrebno imati MS račun. Oba načina rada ovog alata su u potpunosti besplatna, no traži se da korisnik ima elektroničku poštu (engl. *e-mail*), kako bi korištenje alata bilo efikasno.



Slika 14: Power BI data source

Power BI podržava različite formate izvora iz kojih se mogu crpiti podaci i prema tome organizirati izvještaj, kao što je i prikazano na slici broj 14. Slika prikazuje alat gdje se može vidjeti da je to platforma u koju se mogu dovući podaci različitih formata bez obzira na izvor (engl. *Source*) iz kojeg dolaze, Power BI je dovoljno inteligentan da bi uspješno pročitao informacije koje su potrebne za daljnji prikaz podataka.

Power BI pruža veliki spektar mogućnosti obrade i vizualizacije podataka. Može se čak opisati i kao alat koji je relativno jednostavan za korištenje, „S Power BI alatom korisnik ima mogućnost stvoriti analizu podataka postavljajući alatu pitanja putem prirodnog jezika, dakle, nekodiranom engleskom jeziku bez posebne sintakse.“ (Ferrari, Russo, Introducing Microsoft Power BI). Autori su u radu napomenuli da se može koristiti čak i prirodni jezik (engl. *Natural-language queries*) za postavljanje upita, a ta mogućnost može biti itekako privlačna korisnicima i naravno puno pristupačniji način izrade izvještaja.



Slika 15: Power BI – Natural language queries

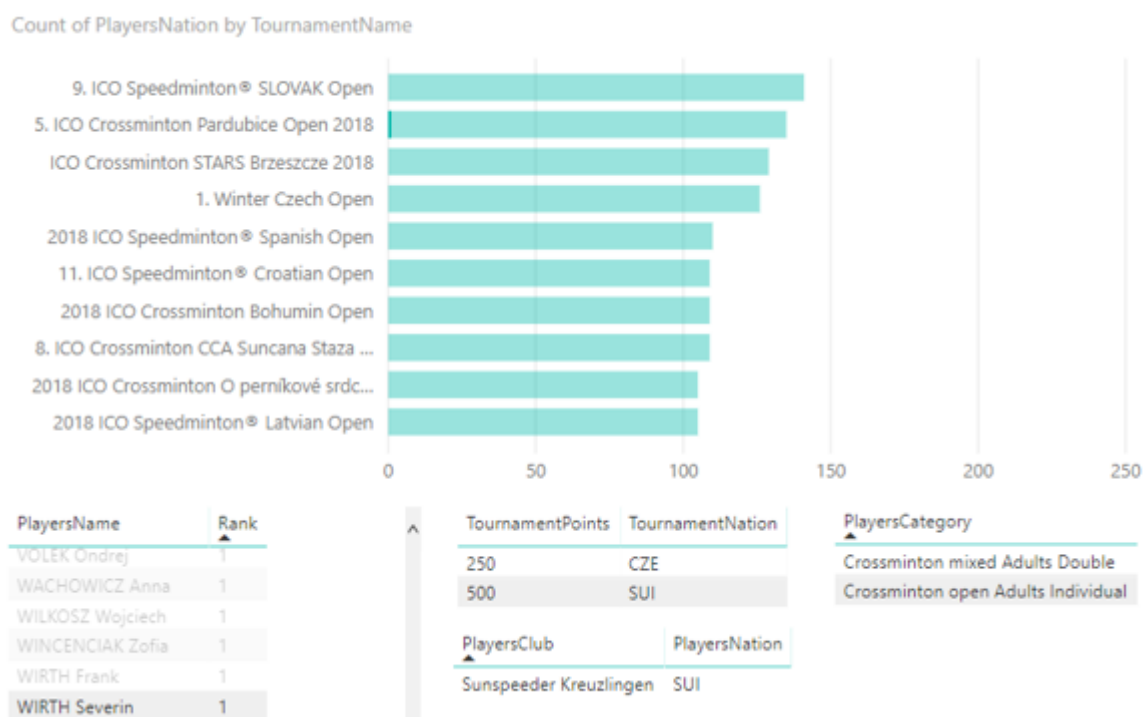
Na slici broj 15. može se vidjeti upotreba mogućnosti korištenja prirodnog jezika za postavljanje upita. U praksi je postavljen upit „show drzava“ nakon čega Power BI ispisuje sve države koje nalazi kao države gdje su odigrani turniri u sezoni. Kako bi se moglo dokazati koliko je inteligentan ovaj alat, postavljen je i drugi upit koji glasi „count turnir drzava“, a tu Power BI uočava da postoji više mogućnosti, „turnir drzava“ te „igraci drzava“. Kako bi se pokazalo da se radi o točnom izračunu samog alata može se izračunata brojka od 17 država, usporediti sa samim popisom koji je prisutan s lijeve strane slike broj 15 gdje je sastavljen čitavi popis svih 17 iskalkuliranih država čime se i potvrđuje ispravnost izračuna. Kako bi se modificirao izvještaj, potrebno je uključiti nekoliko vizualnih figura i to zajedno povezati. Prema autorima Ferrari i Russo (2016.), izvještaj je zapravo samo skup podataka koji su vizualno privlačno organizirani kako bi mogli korisnike jednostavnije informirati. Autori napominju da je

jasno kako samo jedna figura nije dovoljna da bi se dobio smisao izvještaja, u konačnici to iziskuje nešto detaljniji opis i primjereniji tip vizualizacije. Osim korištenja naredbi putem prirodnog jezika, Power BI je uspješan u primjeni umjetne inteligencije (engl. *Artificial intelligence*) kako bi podaci bili samostalno analizirani od strane samog programa, odnosno alata, „Alat koristi izrazito sofisticirane algoritme čija brzina izvršavanja ovisi o veličini i kompleksnosti seta podataka.“ (Ferrari, Russo, Introducing Microsoft Power BI). Jedna od izrazito naprednih opcija ovog alata se naziva, takozvani „Brzi pregled“ (engl. *Quick Insights*). *Quick Insights* omogućuje da se izvještaj sam modificira na način da alat pretpostavlja koji bi izvještaji mogli biti zanimljivi korisniku te koji su one korisne tematike, zaključujući na temelju toga što se nalazi u dostupnim mu podacima. Alat naravno ne poznaje o kojoj se temi tu radi, stoga se korisnik ne može oslanjati isključivo na ovaj način rada jer ništa ne može zamijeniti čovjekovo razumijevanje te zahtjeve koje sam treba provesti kako bi dobio željene rezultate. Iz tog razloga potrebno je izrađivati vlastite izvještaje. Primjer takvih izvještaja vidljivi su na slici broj 16 i 17.



Slika 16: Power BI (1. izvještaj)

Slika broj 16 je primjer jednog takvog izvještaja. Napravljen je izvještaj jednog crossminton turnira iz 2018. godine. Gornji dio izvještaja prikazuje popis turnira zajedno s brojem igrača koji su sudjelovali na skali od 0 do 350. Označen turnir nad kojim će biti izvršen ostatak izvještaja stoji pod nazivom „2018 ICO Crossminton Robocom Swiss Open“. Sljedeći podaci koji su predmetom zanimanja čini popis igrača koji su osvojili prvo mjesto na turniru, to je ime i prezime igrača koji nose „Rank“ 1, u polje „Filters“ je uključen isključivo „Rank 1“. Traži se opći podatak o tome koji turnir je izabran i koliko bodova taj turnir nosi. Nadalje su ispisani klubovi koji su sudjelovali na turniru, te konačno odigrane kategorije. Power BI omogućuje da se ide u dubinu samog izvještaja te da se iscrpe informacije dokle god je moguće. Kao što je prikazano na slici broj 17 gdje se može iščitati više podataka, na primjer, o pojedinom igraču.



Slika 17: Power BI (2. Izvještaj)

Na slici broj 17 korišten je isti izvještaj samo su postavljeni drugačiji uvjeti. U ovom primjeru označuje se primarno igrač u polju „PlayersName“ prema kojem se dalje formulira ostatak izvještaja. Odabrani igrač je „WIRTH Severin“ i u nastavku je stavka interesa. Iz razloga jer se i dalje koristi nepromijenjen filter, traže se podaci odabrane osobe s isključivo tom pozicijom, „Rank“ broj 1, dakle osvajači turnira. Za odabranog

igrača se dobije popis turnira koji je taj igrač osvojio te koliko je to bodova bilo. Također ispisuju se osobni podaci kao što je ime kluba i nacija traženog igrača te konačno, koju kategoriju je taj igrač osvojio. Ferrari i Russo (2016.) nalažu da se u izvještaju korisnik može slobodno izražavati s nešto širim spektrom vizualizacija i informacija, kao što je to i učinjeno u primjeru izvještaja na slici 17. Slika broj 17 prikazuje broj osvojenih turnira koji su prikazani na vrhu izvještaja pomoću složenog stupčastog grafa (engl. *Stacked bar chart*), tamo je tamnijom bojom označen osvojen turnir i u podnožju je prisutna skala brojeva. Tim načinom prikaza je izvještaj vizualno privlačniji i jasnije se može iščitati koliko turnira je odabrani igrač osvojio.

4.4. Usporedba alata

Nakon što je većina Microsoft alata opisana i uzeto u obzir koje su sve njihove sposobnosti, može se sastaviti procjena koji bi od njih bili prikladniji za koju vrstu poslova ili podataka te iz kojih je to razloga tako. Prvotno je potrebno odrediti o koliko se to količini podataka radi, ako se govori o ogromnim količinama podataka onda sami MS Excel možda i nije najučinkovitije rješenje, alat ne posjeduje toliko visoku razinu performansi kako bi zadovoljio sve potrebne kriterije. Ukoliko je dovoljan kapacitet koji pruža MS Excel i njegovi dodaci kao što je Power Pivot, tada bi se trebalo uzeti u obzir da bi upravo taj alat došao kao izrazito prikladno rješenje. Kada je riječ o velikim skladištima podataka najbolje je koristiti SQL Server koji pruža široki spektar rješenja za potrebe manevriranja podacima. Microsoft je kompanija s mnoštvom široko poznatih i korištenih alata, no potrebna je licenca za korištenje koja naravno nosi svoju cijenu. Jedan od spomenutih alata za poslovnu inteligenciju je upravo Power BI koji svoja rješenja i tehnologiju pruža u potpunosti besplatno na korištenje. Power BI se također ističe zbog toga što omogućuje pretraživanje i stvaranje izvještaja koristeći samo prirodni jezik, bez potrebe kodiranja, te i konačno automatizacija izvještaja od strane alata koristeći se opcijom *Quick Insights*. Stoga za vizualizaciju podataka Power BI danas stoji kao jedno od najboljih rješenja na tržištu uz spomenutu činjenicu da je to besplatan alat, izrazito jednostavan za korištenje (engl. *User friendly*) te dostupan svima.

5. Zaključak

Nakon prolaska kroz pojam poslovne inteligencije, njezine važnosti u svakodnevnom poslovanju te razrade bitnih alata za uspostavljanje takvog oblika poslovanja može se zaključiti da je to neophodan aspekt svakog uspješnog poslovanja. U samoj pozadini nečega što će u konačnici pojedinac ili čitava organizacija imati kao uvid u svoje djelovanje potrebno je odraditi iscrpan posao glede upravljanja informacijama. To je činjenica koja upućuje koliko je proces analize informacija neizostavan u budućnosti poslovanja. Nakon što su informacije zaprimljene, njih je potrebno pohraniti uz proces sortiranja kako baze podataka ne bi rezultirale prekomjernim nakupljanjem smeća (engl. *Junk*) koja će samo stvoriti nered u nečemu što bi trebao biti uređeni sustav informacija. Nakon što su sortirane, potrebno je odrediti primjereni alat za njihovo manipuliranje. Najbitniji dio tog postupka je konačni prikaz i rezultat koji se želi postići. Taj rezultat je najbolje vidljiv kada je prikazan nekim vizualnim alatom. Vizualizacija je zapravo proces pretvaranja skupa informacija u jednostavniji i jasniji prikaz kojim se želi nešto poručiti. Proces koji omogućuje korisnicima da vide i razumiju podatke koji im se žele prikazati. Naveden je kao jedan od alata koji se probija na tržište velikom brzinom te uzima status vodećeg u vizualnom prikazu podataka, a tu ulogu nosi upravo Power BI. Power BI se našao kao alat koji je prikladan za izradu poslovne analize, dijeljenje vizualizacija te automatsko generiranje promjena, ukoliko one nastupaju. Temeljem toga nije čudno da se upravo Power BI postavlja kao najpristupačniji alat koji se trenutno nalazi na tržištu i svakodnevno raste po pitanju korisnika.

6. Literatura

1. Carić T. i Buntić M. (2015). *Uvod u relacijske baze podataka*
2. Covington D. (2016). *Analytics (Data Science, Data Analysis, and Predictive Analytics for Business)*
3. Ćurko K. *Skladište podataka – Sustav za potporu odlučivanju*
4. DW4U. Darawarehouse4u.info. URL: https://www.datawarehouse4u.info/index_en.html (2019-05-27)
5. Ferrari A., Russo M. (2016). *Introducing Microsoft Power BI*
6. GDP informer. URL: <https://gdprinformer.com/hr/vodic-kroz-gdpr> (2019-05-17)
7. Gornik D. (2002). *Data Modeling for Data Warehouses*
8. Lauer J. et al (2010). *How to Choose the Right Business Intelligence Technology to Suit Your Style*
9. Kimball R., Ross M. (2002). *The Data Warehouse Toolkit* (2019-05-20)
10. Microsoft. URL: <https://www.microsoft.com/hr-hr/> (2019-05-20)
11. Powell B. (2017). *Microsoft Power BI Cookbook*
12. Search Business Analytics. URL: <https://searchbusinessanalytics.techtarget.com/definition/business-intelligence-BI> (2019-05-02)
13. SQLBI. URL: <https://www.sqlbi.com/> (2019-05-27)
14. Tomaš A. (2014). *Poslovna inteligencija – oblikovanje OLAP kocke na primjeru. Završni rad. Varaždin: Fakultet organizacije i informatike*
15. Tutorialspoint. URL: https://www.tutorialspoint.com/power_bi/power_bi_introduction.htm (2019-05-02)
16. Withee K. (2010). *Microsoft Business Intelligence for Dummies*

7. Popis slika

Slika 1: Primjer integracije podataka	7
Slika 2: Razlika između transakcijskog sustava i skladišta podataka	8
Slika 3: Činjenična tablica	11
Slika 4: Funkcija dimenzijske i činjenične tablice	12
Slika 5: Primjer zvjezdastog modela (engl. <i>Star schema</i>)	13
Slika 6: Pahuljasta shema (engl. <i>Snowflake schema</i>)	14
Slika 7: Tablica u SQL-u.....	15
Slika 8: SQL arhitektura	17
Slika 9: MS Excel primjer jednostavne baze igrača.....	19
Slika 10: MS Excel prikaz mogućnosti automatizacije zapisa	19
Slika 11: MS Excel Pivot tablica	20
Slika 12: Power Pivot	22
Slika 13: SSIS (engl. <i>SQL Server Integration Server</i>)	25
Slika 14: Power BI data source	27
Slika 15: Power BI – Natural language queries	28
Slika 16: Power BI (1. izvještaj).....	29
Slika 17: Power BI (2. Izvještaj).....	30

Alati za poslovnu inteligenciju

Sažetak

U radu je definirana i opisana poslovna inteligencija, te je prikazana njena važnost, potreba i uloga u svakodnevnom poslovanju. Opisani su alati koji se koriste u okruženju poslovne inteligencije, a poseban fokus je na Microsoft alatima koji su danas najšire dostupni i izrazito razvijeni. Analiziran je način na koji alati funkcioniraju, prikazan je pregled njihovih mogućnosti i karakteristika. Glavni cilj rada je prikazati nužnost vizualizacije podataka koja se može napraviti i organizirati na različite načine kako bi podaci bili upotrebljivi korisnicima. Zbog toga je poseban naglasak stavljen na Microsoftov Power BI alat i njegovu djelotvornost u modernom poslovanju.

Ključne riječi: poslovna inteligencija, Microsoft alati, Power BI, poslovna analiza, vizualizacija

Tools for business intelligence

Summary

In this paperwork the term of business intelligence is defined, also it's importance, need and role in everyday business is showed. The tools which are used in business intelligence environment are described, and the focus is on Microsoft tools which are today widely available as well as extremely developed. Furthermore, the way that the tools are functioning are analyzed and the review of their possibilities and characteristics have been showed. The main aim of this paperwork is to show the necessity of data visualization which can be made and organized on multiple ways to make the data effective for users. Because of it, the main accent is on the Microsoft Power BI tool and it's efficiently in modern business.

Key words: business intelligence, Microsoft tools, Power BI, Business analysis, visualization